Three dimensional texture analysis with machine learning provides incremental predictive information for successful shock wave lithotripsy in patients with kidney stones

Mannil, Manoj ; von Spiczak, Jochen ; Hermanns, Thomas ; Poyet, Cédric ; Alkadhi, Hatem ; Fankhauser, Christian Daniel

Abstract: PURPOSE To determine the predictive value of three-dimensional texture analysis (3D-TA) in computed tomography (CT) images for successful shock wave lithotripsy (SWL) in patients with kidney stones. MATERIAL AND METHODS Patients with pre and postoperative CT scans, previously untreated kidney stones and a stone diameter of 5-20 mm were included. A total of 224 3D-TA features of each kidney stone, including the attenuation measured in Hounsfield Units (HU), and the clinical variables body mass index (BMI), initial stone size, and skin-to-stone distance (SSD) were analyzed using five commonly used machine learning models. The data set was split in a ratio of 2/3 for model derivation and 1/3 for validation. Machine learning-based predictions for SWL success in the validation cohort were evaluated calculating sensitivity, specificity, and the area-under-the-curve (AUC). RESULTS For SWL success the three clinical variables BMI, initial stone size and SSD showed AUCs of 0.68, 0.58 and 0.63 respectively, but no predictive value for HU was found. A RandomForest classifier using three 3D-TA features had an AUC of 0.79. By combining these three 3D-TA features with clinical variables, the discriminatory accuracy improved further with an AUC of 0.85 for 3D-TA features and SSD, an AUC of 0.8 for 3D-TA features and BMI and an AUC of 0.81 for 3D-TA and stone size. CONCLUSION This preliminary study indicates that the clinical variables BMI, initial stone size and SSD show limited value for predicting SWL success, while the HU values of stones were not predictive. Selected 3D-TA features identified by machine learning provided incremental accuracy for predicting the success to SWL.

DOI: https://doi.org/10.1016/j.juro.2018.04.059

Originally published at:
Mannil, Manoj; von Spiczak, Jochen; Hermanns, Thomas; Poyet, Cédric; Alkadhi, Hatem; Fankhauser, Christian Daniel (2018). Three dimensional texture analysis with machine learning provides incremental

predictive information for successful shock wave lithotripsy in patients with kidney stones. Journal of Urology, 200(4):829-836.
DOI: https://doi.org/10.1016/j.juro.2018.04.059

Author's Accepted Manuscript

Three dimensional texture analysis with machine learning provides incremental predictive information for successful shock wave lithotripsy in patients with kidney stones

Manoj Mannil , Jochen von Spiczak , Thomas Hermanns , Cédric Poyet , Hatem Alkadhi , Christian Daniel Fankhauser

PII: S0022-5347(18)42986-2
DOI: 10.1016/j.juro.2018.04.059
Reference: JURO 15564

To appear in: *The Journal of Urology*
Accepted Date: 8 April 2018

Please cite this article as: Mannil M, von Spiczak J, Hermanns T, Poyet C, Alkadhi H, Fankhauser CD, Three dimensional texture analysis with machine learning provides incremental predictive information for successful shock wave lithotripsy in patients with kidney stones, *The Journal of Urology®* (2018), doi: 10.1016/j.juro.2018.04.059.

Three dimensional texture analysis with machine learning provides incremental predictive information for successful shock wave lithotripsy in patients with kidney stones

Manoj Mannil[1], Jochen von Spiczak[1], Thomas Hermanns[2], Cédric Poyet[2], Hatem Alkadhi[1], Christian Daniel Fankhauser[2*]

[1] Institute of Diagnostic and Interventional Radiology, University Hospital Zurich, University of Zurich, Switzerland

[2] Department of Urology, University Hospital Zurich, University of Zurich, Switzerland

Corresponding author

Manoj Mannil[1], manoj.mannil@usz.ch, Institute of Diagnostic and Interventional Radiology, University Hospital Zurich, University of Zurich, Switzerland

Jochen von Spiczak[1], jochen.vonspiczak@usz.ch, Institute of Diagnostic and Interventional Radiology, University Hospital Zurich, University of Zurich, Switzerland

Thomas Hermanns[2], thomas.hermanns@usz.ch, Department of Urology, University Hospital Zurich, University of Zurich, Switzerland

Cédric Poyet[2], cedric.poyet@usz.ch, Department of Urology, University Hospital Zurich, University of Zurich, Switzerland

Hatem Alkadhi[1], hatem.alkadhi@usz.ch, Institute of Diagnostic and Interventional Radiology, University Hospital Zurich

Fankhauser Christian Daniel[2*], christian.fankhauser@usz.ch, Department of Urology, University Hospital, University of Zurich, Frauenklinikstrasse 10, 8091 Zurich, Switzerland. Phone: +41 44 255 54 40; Fax: +41 255 54 55;

Running title: Texture analysis predicts SWL success

KEYWORDS: Kidney Calculi; Lithotripsy; Treatment Outcome

Word count: 2497 words

Abstract

Purpose: To determine the predictive value of three-dimensional texture analysis (3D-TA) in computed tomography (CT) images for successful shock wave lithotripsy (SWL) in patients with kidney stones.

Material and methods: Patients with pre and postoperative CT scans, previously untreated kidney stones and a stone diameter of 5-20 mm were included. A total of 224 3D-TA features of each kidney stone, including the attenuation measured in Hounsfield Units (HU), and the clinical variables body mass index (BMI), initial stone size, and skin-to-stone distance (SSD) were analyzed using five commonly used machine learning models. The data set was split in a ratio of 2/3 for model derivation and 1/3 for validation. Machine learning-based predictions for SWL success in the validation cohort were evaluated calculating sensitivity, specificity, and the area-under-the-curve (AUC).

Results: For SWL success the three clinical variables BMI, initial stone size and SSD showed AUCs of 0.68, 0.58 and 0.63 respectively, but no predictive value for HU was found. A RandomForest classifier using three 3D-TA features had an AUC of 0.79. By combining these three 3D-TA features with clinical variables, the discriminatory accuracy improved further with an AUC of 0.85 for 3D-TA features and SSD, an AUC of 0.8 for 3D-TA features and BMI and an AUC of 0.81 for 3D-TA and stone size.

Conclusion: This preliminary study indicates that the clinical variables BMI, initial stone size and SSD show limited value for predicting SWL success, while the HU values of stones were not predictive. Selected 3D-TA features identified by machine learning provided incremental accuracy for predicting the success to SWL.

Introduction

The recommended treatment options for kidney stones <20 mm are extracorporeal shock wave lithotripsy (SWL) or flexible ureterorenoscopy (URS). SWL is convenient for patients because it can be performed in an outpatient setting without the need of urogenital access. However, the number of shock waves which can safely be applied are limited and stone free rates of as low as 21-25% [1-3] for lower pole and 40-85% for non-lower pole stones [2-5] have been reported. Therefore, predictive variables for successful first line SWL would allow to stratify patients to either undergo primary SWL or URS.

Several variables including the stone size [6], body mass index (BMI) [7], skin-to-stone distance (SSD) [8, 9], anatomical factors [10, 11], stone composition [9] and computed tomography (CT) attenuation values measured in Hounsfield Units (HU) [12-15] have been reported as potential predictors for SWL success. Contradictory results regarding the association and clinically useful cut-off values for HU have been published. So far, it remains difficult to predict which patients will have successful primary SWL.

Texture analysis (TA) detects distinct quantifiable differences of stone characteristics that may not be seen in a purely visual analysis. Our group [16] and Cui et al. [17] have recently demonstrated that two dimensional (2D) TA can predict successful SWL in an in-vitro stone model. Latest developments in software technology have enabled three dimensional (3D) TA instead of only 2D-TA analysis. The aim of this study was to apply 3D-TA to an in-vitro stone model and to explore the predictive information of 3D-TA for SWL success in an in-vivo cohort of patients with kidney stones.

Materials and Methods

Patient population

We retrospectively identified patients with kidney stones treated with SWL at our tertiary care center between 2003-2016. Patients with previously untreated kidney stones and a stone diameter of 5-20mm were included. Informed consent was waived for this retrospective cohort study by the local ethic committee (STV KEK-ZH 2014-0198). SWL was performed as previously described [3]. Successful SWL stone disintegration was defined as no residual stones or residual stone diameter <2mm on CT[18].

CT data acquisition and analysis

Over the inclusion period, non-contrast enhanced abdominal CT images were acquired on eight different CT scanners of three different vendors. Detailed scan parameters are summarized in **Supplementary Table 1**. One blinded reader (MM, with 4 years of experience in abdominal radiology) using the workstation (BARCO display Nio Color 3MP MDNC-3421, Kortrijk, Belgium) and the picture archiving and communication system (Impax 6, Agfa Healthcare, Mortsel, Belgium) documented the following variables: stone size before SWL, total number of stones, number of stones per patient, location of stones, SSD, and residual stone size after SWL.

All measurements were performed by a single reader (with 5 years experience in radiology) on a picture archiving and communication system (PACS) workstation. Stone location was assessed in multiplanar images. As previously described [19] CT attenuation in HU and stone size were measured in a standard bone window (width/level -1,120/-300). Images with largest stone diameter were selected to define maximum stone size. An ellipsoid region-of-interest (ROI), slightly smaller in size than the stone in magnified images, was used to measure CT attenuation values. Skin-to-stone distance was measured with radiographic calipers. The measurements were performed in axial images from the point of the largest stone diameter at a set angle from the horizontal line. The mean of measurements at 0°, 45°, and 90° was calculated as previously described [19].

Texture analysis

On all CT data, gray level normalization was performed using the "1-99%" method for correcting small technical intra- and interscanner variations [20].

As several 3D-TA features require identical spatial resolution to be comparable, pixel spacing was normalized to $0.4 \times 0.4 mm^2$ using an in-house MATLAB script (The MathWorks Inc., Natick, Massachusetts, USA).

3D-TA was performed in all stones by one blinded reader (MM) using freely available software (MaZda, version 4.6, Institute of Electronics, Technical University of Lodz, Lodz, Poland). Polygonal 2D regions-of-interest (ROIs) were drawn on a stack of DICOM images resulting in one volume-of-interest (VOI) per stone. VOI delineation was restricted to those urinary stones that underwent subsequent SWL; surrounding structures were carefully excluded (**Figure 1**).

For testing intrareader reproducibility of 3D-TA features, VOI delineation was repeated by the same reader after 3 weeks to avoid recall bias. For testing interreader reliability, VOI delineation was performed by a second blinded reader (with 2 years experience in radiology). Overall, 224 3D-TA features per VOI were computed. The selected 3D-TA features originate from four main categories:

(*i*) Histogram (12 bits-per-pixel): MinNorm3D, MaxNorm3D, Mean3D, Variance3D, Skewness3D, Kurtosis3D, Perc.01%3D, Perc.10%3D, Perc.50%3D, Perc.90%3D, Perc.99%3D;

(*ii*) Grey-Level Co-Occurrence Matrix (GLCM) (6 bits-per-pixel) at three inter-pixel distances: angular second moment, contrast, correlation, entropy, sum entropy, sum of squares, sum average, sum variance, inverse different moment, difference entropy, difference variance;

(*iii*) Run-Length Matrix (RLM) (6 bits-per-pixel) at five angles: 0°, 90°, 135°, 180°, Z: run-length non-uniformity, grey-level non-uniformity, long run emphasis, short run emphasis, fraction of image in runs; and

(*iv*) Absolute gradient (6 bits-per-pixel): gradient mean, variance, skewness, kurtosis, and non-zeros.

TA feature selection and dimension reduction

Feature selection and dimension reduction was performed on the 224 TA features as follows: First, we removed those TA features with reduced intra- and interreader reproducibility. For doing so, intraclass correlation coefficients (ICC) were calculated for each pair of variables. An ICC of 0.61-0.8 were interpreted as substantial and 0.81-1.00 as excellent agreement. We excluded TA features with an ICC ≤0.6 from further analyses. Then, feature reduction and classification analysis were performed similar to Sogawa et al. [21] using open source data mining software (WEKA, version 3.8.0; University of Waikato, Hamilton, New Zealand). The in-built feature selection filter was used on a separate model derivation data set to evaluate the worth of single features (including both TA and non-TA features) by taking into account the individual predictive value for SWL success and the redundancy between highly correlating features.

Statistical analysis

Continuous normally distributed variables are expressed as means±standard deviation and continuous non-normally distributed variables are presented as median±interquartile ranges. Categorical variables were expressed as frequencies or percentages. Normality was assessed using the Shapiro-Wilk test. Grouped differences in initial stone size among training and test groups were tested using the Kruskal-Wallis Test.

After dimension reduction and feature selection, remaining 3D-TA features and three additional features (i.e., BMI, initial stone size, and SSD) were tested in various combinations using five commonly applied machine learning algorithms for classification of SWL success: J48 decision tree, k-nearest neighbor (kNN), artificial neural network (aNN) with backpropagation (Multilayer Perceptron), Random Forest, and sequential minimal optimization (SMO). To account for overfitting, the data set was randomly split in the recommended ratio of 2/3 for model-derivation and 1/3 for validation. Machine learning-based predictions of SWL success on the validation cohort were evaluated by calculating sensitivity, specificity, and area-under-the-curve (AUC) from receiver-operating characteristics (ROC).

Multivariate analysis with a stepwise forward approach was performed for all selected non-TA and TA features by means of binary logistic regression on the entire data set. The resulting odds ratio, 95% confidence-interval and corresponding p-value of parametric Wald test were noted.

Differences in aforementioned test characteristics were tested using paired t-tests with Bonferroni correction for multiple comparisons. Differences in mean numbers of shockwaves and attenuation between both groups were tested using the Mann-Whitney U test. Statistical analyses were conducted using commercially available software (SPSS 23.0; IBM Corp., Armonk, New York, USA). A two-tailed $P<0.05$ was considered statistically significant.

Results

Study population

Of 101 patients, 20 (20%) were excluded because of missing non-contrast CT before SWL, 28 (28%) because of missing non-contrast CT after SWL and 2 (2%) patients because of compromised/non-diagnostic CT image quality (**Figure 2**). Thus, the final study population included 51 patients with a mean age of 55±15 years (**Table 1**). Patients underwent CT a median of 34 days before and 83 days after SWL. Mean preoperative stone size was 10.2±5 mm, mean attenuation was 643±314 HU, and mean SSD was 94.7±26 mm. Stone disintegration was successful in 21/51 patients (41%). Patients in the model derivation and validation cohort showed comparable characteristics for preoperative (*P*=0.231) and postoperative (*P*=0.183) stone size, absolute number of stones per patient (*P*=0.234), attenuation in HU (*P*=0.09), SSD (*P*=0.09), and BMI (*P*=0.231) (**Supplementary Table 2**).

Texture feature selection

After exclusion of 3D-TA features showing a reduced inter- and intrareader variability, 147/224 features (66%) remained **(Supplementary table 3)**. Feature selection revealed three 3D-TA features predicting successful SWL: Histogram Kurtosis3D, GLCM S(2,-2,0) SumEntrp, and GLCM S(0,3,0) DifEntrp. Histogram 3D-TA features are based on a normalized histogram vector p(i), defined as ratio between total number of pixels with gray level i, and total number of pixels, while GLCM TA features quantify joint probability pixel distributions. The mean attenuation of kidney stones in CT (measured in HU) did not predict the success to SWL. It must be noted, however, that the above mentioned TA feature Kurtosis3D is derived from the histogram of the attenuation of the kidney stones, which indicates that certain features of attenuation do have predictive capability.

Classification analyses

For SWL success the three clinical variables BMI, initial stone size and SSD showed AUCs of 0.68, 0.58 and 0.63 respectively (**Supplementary Table 4**) and no predictive information for HU could be noted. The RandomForest classifier using three 3D-TA features yielded an AUC of 0.79 (**Table 2**). By combining 3D-TA features and clinical variables, the accuracy improved to an AUC of 0.85 for 3D-TA features and SSD, 0.8 for 3D-TA features and BMI and 0.81 for 3D-TA and initial stone size (**Table 2, Figure 3**). Any other combination of 3D-TA and non-TA features did not further improve the AUC results listed above.

Discussion

Our preliminary study indicates two important findings: First, 3D-TA provides incremental predictive information on the success of SWL in addition to previously known clinical variables BMI, SSD, and stone size; and second, mean CT attenuation values of kidney stones were not predictive for SWL success.

Similar to attenuation values, which are measured in HU, TA represents a data post-processing tool which discloses quantitative information contained within medical images that can be used for lesion detection and clinical outcome prediction. So far, urogenital applications of TA included an improvement of accuracy for prostate cancer detection in apparent diffusion coefficient maps [22], and a distinction of histological subtypes of renal and adrenal tumors on CT [23]. Cui et al. [17] and our group [16] have recently demonstrated that 2D-TA can predict successful SWL with moderate accuracy in an in-vitro stone model.

In this study we were able to improve previous TA methodology in four important aspects. First, we used 3D instead of 2D-TA analysis. Assessment of the whole stone is crucial since urinary stones show variable density/mineral composition and hence, variable CT attenuation (in HU) [24]. Second, we used as many as 224 TA features in a discovery cohort and were able to validate the results in a validation cohort to limit spurious results. Third, we analysed inter- and intrareader reliability and results indicate relatively high inter- and intrareader reliability of 3D-TA. In contrast to previous 2D-TA studies where up to 70% of features had to be excluded because of low intra- and interreader reproducibility[16], we found less variability for 3D stone delineation with only 33% of TA features being discarded. And fourth we used stateof-the-art artificial intelligence and machine learning tools to select the best texture analysis features based on reliability and predictive information for SWL success. Using this advanced and rigorous TA methodology, we could show for the first time in-vivo that 3D-TA is feasible, reproducible, and predictive for SWL success in kidney stones.

Our analyses revealed that three previously unknown texture features namely Kurtosis3D, SumEntrp and DifEntrp were predictive for SWL success. In brief, the 3D-TA feature Kurtosis3D represents the histogram shape of the CT attenuation (in HU) probability distribution. Lower kurtosis values, as seen in kidney stones with successful SWL, are the result of frequent modestly sized attenuation deviations, possibly representing a more uniform stone composition/ architecture. The two other TA features S(2,-2,0) SumEntrp and S(0,3,0) DifEntrp are entropy measures. Entropy is a measure of randomness. Lower entropy for instance represents a more homogeneous image texture.

Our analyses also revealed that attenuation values were not predictive for SWL success. Previous publications found absent [19], weak [25, 26] or good [13-15] associations between HU values of kidney stones and SWL success. In this study we found no predictive value of mean HU values, however, the closely related histogram TA feature Kurtosis3D showed predictive capabilities. Regarding the divergent results between our and some studies mentioned above regarding the HU values of kidney stones, it should be noted that several previous studies used questionable definitions of treatment success, such as “successful treatment in a maximum of 3 SWL sessions” [14] or “residual stones less than 4mm in size after SWL” [27]. Those two outcome definitions have limitations, and given the fact that even clinically insignificant residual stones affect recurrence rates [28], we used a smaller cut-off value in the present study. However in an in-vitro stone model, our group could recently show a weak correlation between HU units and SWL success, depending on the tube voltage of CT [19]. Beside different tube voltages, the reported inconsistencies regarding the predictive value of HU may also be attributed to differences in CT scan protocol settings, ROI delineation, image magnification, and windowing. Altogether the clinical usefulness and the cut-off value for any TA feature has to be defined in future trials, before 3D-TA software tools should be applied in daily clinical routine.

The following study limitations must be acknowledged. Our study included only 51 patients, which limits generalizability of our results. This holds true also for statistical analyses precluding the application of multiple random splits for reducing overfitting. Because of the observational study design potential selection, misclassification, and information bias may have occurred. Furthermore many patients undergoing SWL were not able to collect passing stone fragments for stone composition analysis, and we were therefore not able to include stone composition in our analysis[9]. However, our study included patients with untreated kidney stones for which stone composition analysis is not preoperatively available anyway. Furthermore, because of the large variability in stone fragility to shock waves even within the same mineral composition, the predictive information of stone composition remains limited but might be assessed by 3D-TA[24]. Also, we were not able to add the parameter stone location as another potentially relevant factor predicting the success of SWL because of the limited sample size, and with some patients having several stones at different locations. In order to limit the chance for spurious results we validated our results in a separate validation cohort. Nevertheless multiple testing may lead to type I errors, which might overestimate the described associations.

Conclusion

Our first results indicate that certain TA features, as identified through machine learning, might have the potential to improve prediction of SWL outcome. Future studies with larger sample sizes are required to exactly determine which TA feature and which clinical variables will be useful in clinical practice.

Acknowledgements

None

Conflicts of interest

The authors declare no conflicts of interest.

References

1. Albala, D. M., Assimos, D. G., Clayman, R. V. et al.: Lower pole I: a prospective randomized trial of extracorporeal shock wave lithotripsy and percutaneous nephrostolithotomy for lower pole nephrolithiasis-initial results. J Urol, **166:** 2072, 2001
2. Geraghty, R., Burr, J., Simmonds, N. et al.: Shock wave lithotripsy outcomes for lower pole and non-lower pole stones from a university teaching hospital: Parallel group comparison during the same time period. In: Urol Ann. India, vol. 7, pp. 46-8, 2015
3. Fankhauser, C. D., Hermanns, T., Lieger, L. et al.: Extracorporeal shock wave lithotripsy versus flexible ureterorenoscopy in the treatment of untreated renal calculi. Clinical Kidney Journal**:** sfx151, 2018
4. Neisius, A., Wollner, J., Thomas, C. et al.: Treatment efficacy and outcomes using a third generation shockwave lithotripter. BJU Int, **112:** 972, 2013
5. Abdel-Khalek, M., Sheir, K. Z., Mokhtar, A. A. et al.: Prediction of success rate after extracorporeal shock-wave lithotripsy of renal stones--a multivariate analysis model. Scand J Urol Nephrol, **38:** 161, 2004
6. Kanao, K., Nakashima, J., Nakagawa, K. et al.: Preoperative nomograms for predicting stone-free rate after extracorporeal shock wave lithotripsy. The Journal of urology, **176:** 1453, 2006
7. Park, H. S., Gong, M. K., Yoon, C. Y. et al.: Computed Tomography-Based Novel Prediction Model for the Outcome of Shockwave Lithotripsy in Proximal Ureteral Stones. J Endourol, **30:** 810, 2016

8. Wiesenthal, J. D., Ghiculete, D., Ray, A. A. et al.: A clinical nomogram to predict the successful shock wave lithotripsy of renal and ureteral calculi. The Journal of urology, **186:** 556, 2011

9. Perks, A. E., Schuler, T. D., Lee, J. et al.: Stone attenuation and skin-to-stone distance on computed tomography predicts for stone fragmentation by shock wave lithotripsy. Urology, **72:** 765, 2008

10. Sumino, Y., Mimata, H., Tasaki, Y. et al.: Predictors of lower pole renal stone clearance after extracorporeal shock wave lithotripsy. The Journal of urology, **168:** 1344, 2002

11. Sorensen, C. M., Chandhoke, P. S.: Is lower pole caliceal anatomy predictive of extracorporeal shock wave lithotripsy success for primary lower pole kidney stones? The Journal of urology, **168:** 2377, 2002

12. Kacker, R., Zhao, L., Macejko, A. et al.: Radiographic parameters on noncontrast computerized tomography predictive of shock wave lithotripsy success. The Journal of urology, **179:** 1866, 2008

13. Pareek, G., Armenakas, N. A., Fracchia, J. A.: Hounsfield units on computerized tomography predict stone-free rates after extracorporeal shock wave lithotripsy. The Journal of urology, **169:** 1679, 2003

14. El-Nahas, A. R., El-Assmy, A. M., Mansour, O. et al.: A Prospective Multivariate Analysis of Factors Predicting Stone Disintegration by Extracorporeal Shock Wave Lithotripsy: The Value of High-Resolution Noncontrast Computed Tomography. European Urology, **51:** 1688, 2007

15. Ng, C. F., Siu, D. Y., Wong, A. et al.: Development of a scoring system from noncontrast computerized tomography measurements to improve the selection

of upper ureteral stone for extracorporeal shock wave lithotripsy. J Urol, **181:** 1151, 2009

16. Mannil, M., von Spiczak, J., Hermanns, T. et al.: Prediction of successful shock wave lithotripsy with CT: a phantom study using texture analysis. Abdominal Radiology**:** 1, 2017
17. Cui, H. W., Devlies, W., Ravenscroft, S. et al.: CT Texture Analysis of Ex Vivo Renal Stones Predicts Ease of Fragmentation with Shockwave Lithotripsy. J Endourol, **31:** 694, 2017
18. Rippel, C. A., Nikkel, L., Lin, Y. K. et al.: Residual fragments following ureteroscopic lithotripsy: incidence and predictors on postoperative computerized tomography. J Urol, **188:** 2246, 2012
19. Mullhaupt, G., Engeler, D. S., Schmid, H. P. et al.: How do stone attenuation and skin-to-stone distance in computed tomography influence the performance of shock wave lithotripsy in ureteral stone disease? BMC Urol, **15:** 72, 2015
20. Collewet, G., Strzelecki, M., Mariette, F.: Influence of MRI acquisition protocols and image intensity normalization methods on texture classification. Magn Reson Imaging, **22:** 81, 2004
21. Sogawa, K., Nodera, H., Takamatsu, N. et al.: Neurogenic and Myogenic Diseases: Quantitative Texture Analysis of Muscle US Data for Differentiation. Radiology**:** 160826, 2017
22. Rozenberg, R., Thornhill, R. E., Flood, T. A. et al.: Whole-Tumor Quantitative Apparent Diffusion Coefficient Histogram and Texture Analysis to Predict Gleason Score Upgrading in Intermediate-Risk 3 + 4 = 7 Prostate Cancer. AJR Am J Roentgenol, **206:** 775, 2016

23. Yu, H., Scalera, J., Khalid, M. et al.: Texture analysis as a radiomic marker for differentiating renal tumors. Abdom Radiol (NY), **42:** 2470, 2017

24. Motley, G., Dalrymple, N., Keesling, C. et al.: Hounsfield unit density in the determination of urinary stone composition. Urology, **58:** 170, 2001

25. El-Nahas, A. R., El-Assmy, A. M., Mansour, O. et al.: A prospective multivariate analysis of factors predicting stone disintegration by extracorporeal shock wave lithotripsy: the value of high-resolution noncontrast computed tomography. Eur Urol, **51:** 1688, 2007

26. Largo, R., Stolzmann, P., Fankhauser, C. D. et al.: Predictive value of low tube voltage and dual-energy CT for successful shock wave lithotripsy: an in vitro study. Urolithiasis, **44:** 271, 2016

27. Tanaka, M., Yokota, E., Toyonaga, Y. et al.: Stone attenuation value and cross-sectional area on computed tomography predict the success of shock wave lithotripsy. Korean J Urol, **54:** 454, 2013

28. Hein, S., Miernik, A., Wilhelm, K. et al.: Endoscopically Determined Stone Clearance Predicts Disease Recurrence Within 5 Years After Retrograde Intrarenal Surgery. J Endourol, **30:** 644, 2016

Legends to illustrations

Figure 1: (A) Non-contrast enhanced abdominal CT of a 76-year-old male patient with a stone in the right lower pole (arrow). (B) Volume-of-Interest (VOI) of a kidney stone in a 3D image stack for 3D texture analysis.

Figure 2: Flowchart of the study

Figure 3: Receiver operating characteristics analysis for successful shockwave lithotripsy in different models of three parameters (BMI, initial stone size, skin-to-stone distance) separately and combined with three selected 3D-TA features (Histogram Kurtosis3D, GLCM S(2,-2,0)SumEntrp, and GLCM S(0,3,0)DifEntrp). 3D-TA features combined with selected BMI [kg/m^2] show a corresponding sensitivity and specificity is 76 % each, while the area-under-the-curve (AUC) is 0.8. Similar receiver operating characteristics analysis for successful shockwave lithotripsy in a model combining three 3D-TA features (Histogram Kurtosis3D, GLCM S(2,-2,0)SumEntrp, and GLCM S(0,3,0)DifEntrp) and initial stone size [mm] were obtained. The corresponding sensitivity and specificity is 76 % each, while the AUC is 0.81.

Table 1: Patient demographics in all patients, in patients with successful and unsuccessful SWL.

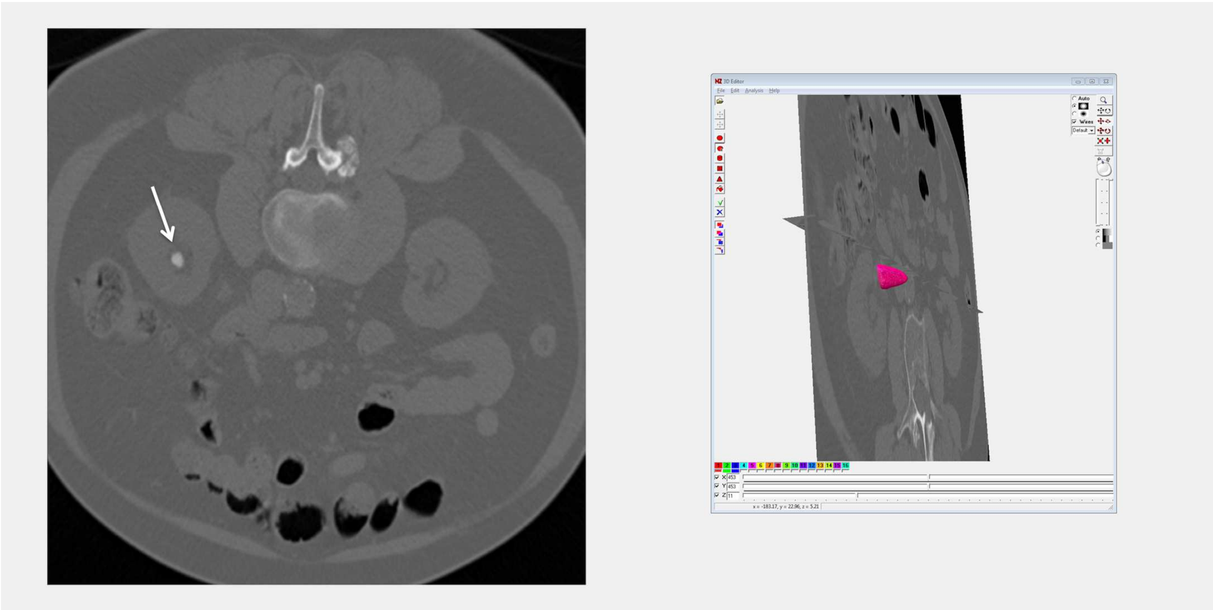
	All	Successful SWL	Unsuccessful SWL
N	51	21 (41%)	30 (59%)
Age [years]	55 ± 15	56 ± 14	54 ± 15
BMI [kg/m^2]	27 ± 5	26 ± 4	28 ± 5
Gender			
Females	14 (27%)	8(38%)	6(20%)
Males	37 (73%)	13(62%)	24(80%)
Stone characteristics			
Renal calculi (n)	3 ± 7	2 ± 1	4 ± 9
Initial calculus size [mm]	10 ± 5	10 ± 5	10 ±5
Calculus size after SWL [mm]	2.3 ± 4	0.05 ± 0.2	6 ± 3
Relative size reduction [%]	66 ± 35	99 ± 4	43 ± 27
Skin-to-stone distance [mm]	95 ± 26	93 ± 20	96 ± 29
Stone localization			
Lower Pole	16 (31%)	7(33%)	9(30%)
Midpole	-	-	-
Upper Pole	4 (8%)	2(10%)	2(7%)
UPR	31 (61%)	12(57%)	19(63%)
Number of shockwaves	2863 ± 388	2952 ± 218	2800 ± 466
Percentage of Energy [%]	70 ± 6	70 ± 5	71 ± 6
Texture features *mean (min; max)*			
Histogram Kurtosis3D	-0.26 ± 1.1 (-1.4;5.1)	-0.56 ± 0.7 (-1.4;0.5)	-0.1 ± 1.3 (-1.3;5.1)
S(2,-2,0) SumEntrp	1.9 ± 0.1 (1.4;2.1)	2 ± 0.1 (1.7;2.1)	1.9 ± 0.1 (1.4;2.1)
S(0,3,0) DifEntrp	1.4 ± 0.1 (1.1;1.6)	1.4 ± 0.1 (1.2;1.5)	1.4 ± 0.1 (1.1;1.6)

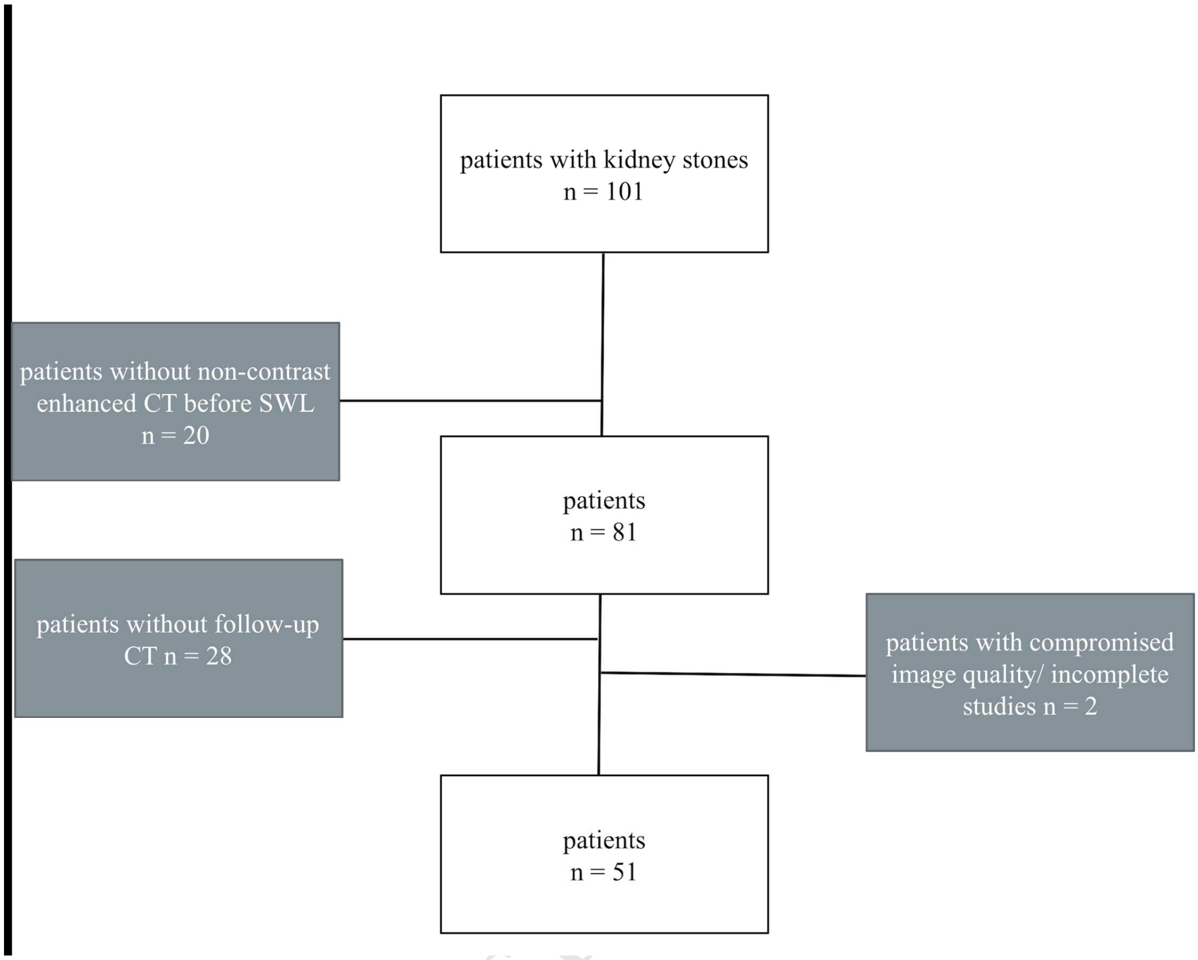
BMI = Body-Mass-Index; max = maximum; min = minimum; SWL = Shockwave lithotripsy; UPR = Ureteropelvic junction

Table 3: Classification results of 3D TA for the prediction of the success to SWL.

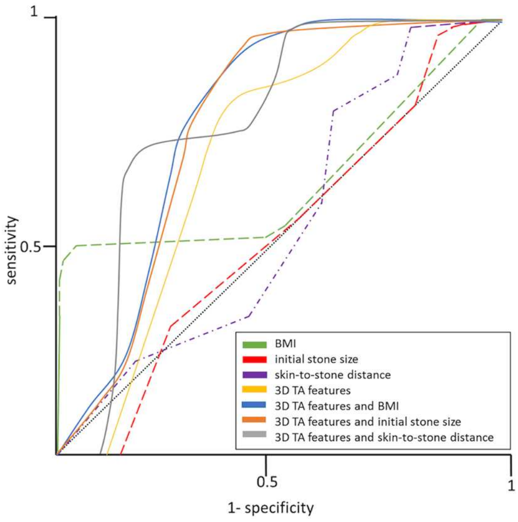
Model	Classifier	Sensitivity	Specificity	AUC
3D TA	J48 (C 4.5)	0.71	0.74	0.72
	kNN	0.53	0.68	0.61
	MultilayerPerceptron (aNN)	0.65	0.72	0.6
	RandomForest	**0.71**	**0.74**	**0.79**
	SMO	0.35	0.63	0.49
3D TA + BMI	J48 (C 4.5)	0.71	0.74	0.72
	kNN	0.47	0.66	0.57
	MultilayerPerceptron (aNN)	0.53	0.51	0.62
	RandomForest	**0.76**	**0.76**	**0.8**
	SMO	0.41	0.65	0.53
3D TA + initial stone size	J48 (C 4.5)	0.71	0.74	0.72
	kNN	0.53	0.51	0.52
	MultilayerPerceptron (aNN)	0.53	0.68	0.6
	RandomForest	**0.76**	**0.76**	**0.81**
	SMO	0.35	0.63	0.49
3D TA + SSD	J48 (C 4.5)	0.59	0.7	0.64
	kNN	0.41	0.65	0.53
	MultilayerPerceptron (aNN)	0.41	0.47	0.46
	RandomForest	**0.65**	**0.72**	**0.85**
	SMO	0.29	0.61	0.45

aNN = artificial neural network; AUC = Area-under-the-curve; DTNB = Decision Table and Naïve Bayes; kNN = k-nearest neighbor; ROC = Receiver-Operating Characteristics; SMO = Sequential minimal optimization; SSD = skin-to-stone distance; SWL = Shock wave lithotripsy; Best classification results printed in **bold**





patients with kidney stones
n = 101
patients without non-contrast enhanced CT before SWL
n = 20
patients
n = 81
patients without follow-up CT n = 28
patients with compromised image quality/ incomplete studies n = 2
patients
n = 51


1
0.5
sensitivity
0.5
1
1- specificity
BMI
initial stone size
skin-to-stone distance
3D TA features
3D TA features and BMI
3D TA features and initial stone size
3D TA features and skin-to-stone distance

Abbreviations

area-under-the-curve (AUC)

body mass index (BMI)

computed tomography (CT)

confidence intervals (CI)

extracorporeal shock wave lithotripsy (SWL)

Hounsfield Units (HU)

odds ratios (ORs)

receiver operating characteristics (ROC)

interquartile ranges (IQR)

standard deviation (SD)

skin to stone distance (SSD)

two-dimensional (2D)

three-dimensional (3D)

texture analysis (TA)